

GMCM “华为杯”第十七届中国研究生数学建模竞赛

中国研究生创新实践系列大赛
“华为杯”第十七届中国研究生
数学建模竞赛

学 校

西安电子科技大学

参赛队号

20107010019

队员姓名

- | | |
|----|-----|
| 1. | 吴涛 |
| 2. | 刘文慧 |
| 3. | 邢卓林 |

中国研究生创新实践系列大赛

“华为杯”第十七届中国研究

数学建模竞赛

题 目

无人机集群协同对抗

摘 要

随着无人机智能自主化的发展，无人机集群作战技术已成为智能作战的重要形式。本文通过分析红、蓝双方的无人机在平面区域上的协同对抗问题，针对不同的作战场景和兵力部署建立相应的数学模型，设计最优拦截策略和突防策略。

针对问题一，我们依据红、蓝双方的兵力设置情况以及在平面上的相对位置关系，构建出基于微分方程的路径规划模型，模拟出对抗双方进行突防和拦截的飞行轨迹，在矩形区域 ABCD 内存在一个可行域，无论红方无人机采取什么样的追击策略，蓝方无人机总能采用合适的策略躲避红方的拦截，并讨论了蓝方的最优突防策略。

针对问题二，我们通过微分博弈方法，以最小通道带宽和最短突防时间为目标建立优化模型，构造两点边值问题来证明并求解无人机的通道带宽下限 M_{min} ，当实际通道带宽大于下限值时，蓝方无人机一定能够突破红方无人机的拦截。并在此情况下，求得蓝方无人机的时间最短突防策略。

针对问题三，我们在问题一、二的模型基础上，合理假设了红方两架运载机的初始位置，每个波次发射 5 架无人机。同时确定了蓝色无人机距离红方无人机集群 15km 时，运载机第二波次发射，且新的无人机集群的运行方向与第一波次发射的无人机集群的运行方向相同，两拨无人机间距在满足条件的基础上，在合理的区间内与当前位置到边界的最短距离成正相关。通过微分博弈方法，以最小通道带宽为目标建立优化模型，证明存在并得到了通道带宽上限。

针对问题四，我们建立了基于 BP 神经网络的多机目标分配模型以及基于强化学习的协同博弈模型，求解出红方采取“1-3-1”模式分配无人机时能有效拦截蓝方，各集群在第一波次发射的无人机架数为(3)-(3,5,3)-(3)，第二波次发射的无人机架数为(7)-(7,5,7)-(7)。然后利用强化学习模型 G2ANet 分析红、蓝双方的对抗方案，发现在博弈均衡条件下有最优拦截策略和最优防守策略。

综上所述，本文通过建立基于微分方程的路径规划模型及博弈模型，分析并求解蓝方无人机在不同场景下的最优突防策略，最后根据 BP 神经网络的训练学习结果，分析出红方无人机的最佳兵力分配方案，同时也提出使用强化学习寻找红方最优拦截策略、蓝方最优突防策略的构想，这对实际的双方对抗战事具有一定参考价值。

关键词：微分法路径规划模型 微分博弈 BP 神经网络 博弈均衡 强化学习

目录

一、 问题重述.....	3
1.1 背景资料与条件.....	3
1.2 需要解决的问题.....	3
二、 问题分析.....	3
2.1 问题一的分析.....	3
2.2 问题二的分析.....	3
2.3 问题三的分析.....	4
2.4 问题四的分析.....	4
三、 模型假设.....	4
四、 符号说明.....	4
五、 模型的建立与求解.....	5
5.1 问题一：微分法路径规划模型.....	5
5.1.1 无人机追逃策略分析.....	5
5.1.2 模型建立.....	6
5.1.3 模型求解.....	9
5.2 问题二：微分博弈模型.....	12
5.2.1 模型建立.....	12
5.2.2 模型求解.....	13
5.2.3 实验结果分析.....	15
5.3 问题三：改进的微分博弈模型.....	16
5.3.1 模型建立.....	16
5.3.2 模型求解.....	17
5.3.3 实验结果分析.....	18
5.4 基于 BP 神经网络的目标分配及博弈模型.....	19
5.4.1 方法分析.....	19
1、多机目标分配问题.....	19
2、双方协同博弈问题.....	19
5.4.2 基于 BP 神经网络的目标分配模型.....	20
1、模型建立.....	20
2、模型求解.....	22
3、仿真结果.....	23
5.4.3 基于强化学习的协同博弈模型.....	24
六、 模型的评价.....	25
6.1 模型的优点.....	25
6.2 模型的缺点.....	25
七、 模型的改进与推广.....	25
7.1 模型的改进.....	25
7.2 模型的推广.....	25
参考文献.....	26

一、问题重述

1.1 背景资料与条件

新一代人工智能技术和自主技术快速走向战场，将催生新型作战力量，颠覆传统战争模式，未来战争必将是智能化战争。无人机集群作战作为智能作战的重要形式，正在崭露头角。通过多架无人机协同侦察、协同探测、协同跟踪、协同攻击、协同拦截等，共同完成较复杂的作战任务。

1.2 需要解决的问题

(1) 建模分析蓝方无人机处于矩形区域 $ABCD$ 内哪些位置时，无论红方无人机采用什么样的追击策略，蓝方无人机总能采用合适的策略以躲避红方的拦截，实现成功突防；讨论蓝方无人机相应的最优突防策略。

(2) 建模分析是否存在一个通道带宽 M 的下限 M_{min} ，当实际通道带宽 M 比 M_{min} 大时，蓝方无人机一定能突破红方无人机集群的拦截；给出此种情形下蓝方无人机时间最短的突防策略。

(3) 讨论红方两架运载机两个波次发射的无人机数量、每架运载机第二波次发射的时刻和位置以及第二波次发射的无人机集群的中心位置，以实现最优的拦截效果；进一步具体建模分析是否存在一个通道带宽 M 的上限 M_{max} ，当实际通道带宽 M 小于 M_{max} 时，无论蓝方无人机采用什么样的突防策略，红方无人机集群均存在相应的拦截策略，在区域 $ABCD$ 内成功阻止蓝方无人机的突防。

(4) 讨论红方兵力部署及最优拦截策略和蓝方最优突防策略。

二、问题分析

2.1 问题一的分析

根据红方无人机的初始位置部署以及红、蓝双方无人机飞行速度和转弯半径约束，寻找蓝方在指定 $ABCD$ 矩形区域内一定可以成功突防的可行位置，以及相应的最优突防策略。我们建立微分法路径规划模型^[1]，借助红蓝双方追逃过程中的位置关系构建微分方程，规划出两者的飞行轨迹进行问题的分析求解。

2.2 问题二的分析

我们首先要根据需求合理确定 2 个无人机集群圆周中心的初始位置，然后根据微分博弈^{[2][3]}的方法建立模型来寻找鞍点来证明存在一个通道带宽下限，使得当实际通道带宽大于该带宽下限时，蓝方无人机一定能突破红方无人机集群的拦截。同时通过求解模型得到该鞍点对应的策略，即可得到蓝方无人机时间最短的突防策略。

2.3 问题三的分析

首先通过分析讨论合理地部署红方两架运载机两个波次发射的无人机数量、每架运载机第二波次发射的时刻和位置以及第二波次发射的无人机集群的中心位置。然后我们继续采用微分博弈方法来证明求解通道带宽上限，使得实际通道带宽小于该带宽上限时，无论蓝方无人机采用什么样的突防策略，红方无人机集群均存在相应的拦截策略能成功阻止蓝方无人机的突防。

2.4 问题四的分析

当红、蓝双方的无人机数量都显著增加时，战场态势呈现为多智能体对抗的场景。为了成功拦截蓝方无人机，红方需要对其 50 架无人机进行有效部署。在最佳兵力分配方案的基础上，才能进一步分析两方作战策略，而双方对抗时必须不断学习，最终要满足博弈均衡条件才有相对而言的最优策略。

针对红方兵力分配问题，由于缺乏大量先验信息，可以通过训练 BP 神经网络^{[4]、[5]}学习出拦截效果最好的分配方案。针对红蓝双方博弈问题，可以使用强化学习方法学习到使得己方收益最大的飞行轨迹，这种最优拦截或者突防策略要在博弈均衡的条件下确定。

三、模型假设

- (1) 假设红蓝双方的无人机均在同一高度上机动飞行。
- (2) 假设双方无人机能够准确探测对方无人机和己方无人机的实时位置。
- (3) 假设红方无人机集群在拦截蓝方无人机之前能够调整到最佳拦截状态。
- (4) 假设运动速度恒定，方向可以改变。
- (5) 假设蓝方无人机不会主动攻击红方运载机和无人机集群。
- (6) 假设红方运载机能一直与蓝方无人机保持安全距离。
- (7) 假设红蓝双方无人机能够根据对方状态预测对方下一个时刻的状态。
- (8) 假设无人机相邻状态的时间间隔 $\Delta t = 0.1$ 秒。
- (9) 不考虑运载机部署无人机集群需要的时间。

四、符号说明

符号	说明
θ_E	蓝方无人机转弯的偏转角度
θ_{Emin}	蓝方无人机转弯的最小偏转角度
θ_{Emax}	蓝方无人机转弯的最大偏转角度
θ_P	红方无人机转弯的偏转角度
θ_{Pmin}	红方无人机转弯的最小偏转角度

θ_{Pmax}	红方无人机转弯的最大偏转角度
R_E	蓝方无人机转弯的圆周半径
R_{Emin}	蓝方无人机转弯的最小圆周半径
R_P	红方无人机转弯的圆周半径
R_{Pmin}	红方无人机转弯时的最小圆周半径
V_E	蓝方无人机的速度
V_P	红方无人机的速度
M	通道带宽
M_{min}	蓝方无人机一定能突防的通道带宽下限
M_{max}	红方无人机集群一定能拦截的通道带宽上限

五、模型的建立与求解

5.1 问题一：微分法路径规划模型

5.1.1 无人机追逃策略分析

依据任务要求，蓝方无人机要在 360s 内越过边界 CD 即认为突防成功。经过计算发现蓝方无人机从边界 AB 水平直线飞至边界 CD 只需 200s，那么其在矩形区域内采取曲线飞行形式也可以突防成功；但是曲线轨迹变化情况复杂不利于模型建立与求解，我们以突防耗时最短为原则规定蓝方无人机保持水平直线飞行，只有在遇到红方无人机拦截时才进行转弯躲避，而红方无人机可通过探测装置对蓝方无人机进行实时锁定追踪。另外，无人机集群的圆周半径为 100m，相对于整个 50×70 km² 矩形区域内非常小，因此可将蓝方无人机和两个红方无人机集群都近似看作小圆点。

本节基于这一合理的规定，挖掘红、蓝双方无人机在执行拦截与突防任务时的每一时刻的相对位置关系，发现可以建立微分方程模拟出双方的飞行轨迹线，红方在实时锁定蓝方的过程中的飞行路径呈现曲线形式。以 B 点为原点，AB 所在直线为 y 轴，BC 所在直线为 x 轴建立二维直角坐标系，如下图 1 所示。

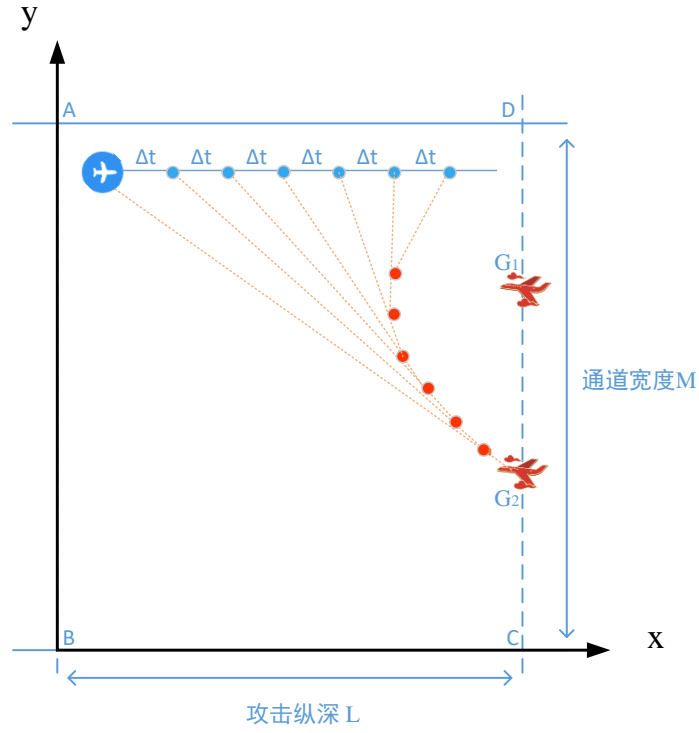


图 1 红、蓝双方轨迹线路图

假定在时刻 t 时，蓝方无人机的位置为 $(x_b(t), y_b(t))$ ，红方无人机集群 FY01 的位置为 $(x_{r01}(t), y_{r01}(t))$ ，红方无人机集群 FY02 的位置为 $(x_{r02}(t), y_{r02}(t))$ 。其中经过很短的一个时间间隔 Δt 后，蓝方无人机水平前进距离为 $v_E \cdot \Delta t$ ，即

$$\begin{cases} x_b(t + \Delta t) = x_b(t) + v_E \cdot \Delta t \\ y_b(t + \Delta t) = y_b(t) \end{cases} \quad (1)$$

红方无人机集群 FY01 水平向左移动 $v_p \cdot \Delta t \cdot \cos \theta_1$ ，竖直向上移动 $v_p \cdot \Delta t \cdot \sin \theta_1$ ，即

$$\begin{cases} x_{r01}(t + \Delta t) = x_{r01}(t) - v_p \cdot \Delta t \cdot \cos \theta_1 \\ y_{r01}(t + \Delta t) = y_{r01}(t) + v_p \cdot \Delta t \cdot \sin \theta_1 \end{cases} \quad (2)$$

红方无人机集群 FY02 水平向左移动 $v_p \cdot \Delta t \cdot \cos \theta_2$ ，竖直向上移动 $v_p \cdot \Delta t \cdot \sin \theta_2$ ，即

$$\begin{cases} x_{r02}(t + \Delta t) = x_{r02}(t) - v_p \cdot \Delta t \cdot \cos \theta_2 \\ y_{r02}(t + \Delta t) = y_{r02}(t) + v_p \cdot \Delta t \cdot \sin \theta_2 \end{cases} \quad (3)$$

依此关系可以确定红方无人机集群的飞行轨迹是一段连续的曲线。

5.1.2 模型建立

1、模型介绍

依据无人机的飞行速度和时间间隔，以及红方无人机集群对蓝方无人机的实时目标锁

定策略，可以确定相邻两个时刻点的无人机位置。当这个时间间隔非常小时，可以分别对红、蓝双方无人机拟合出一条连续的飞行轨迹线。

2、参数说明

表 1 给定参数表

L	M	v_E	v_P	R_E	R_P	θ_1	θ_2
50km	70km	250m/s	200m/s	500m	350m	rl-b	rr-b

在表 1 中， θ_1 表示为红方无人机集群 FY01 与蓝方无人机连线与水平方向的夹角， θ_2 表示红方无人机集群 FY02 与蓝方无人机连线与水平方向的夹角，如下图 2 所示。

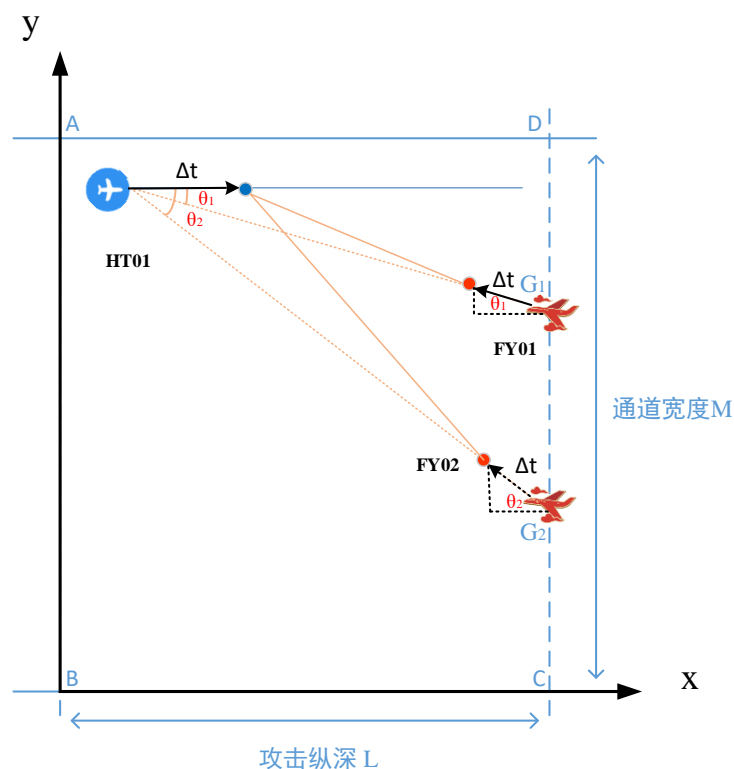


图 2 相邻两时刻无人机位置图

3、场景分析

题目要求无论红方采取什么策略时，蓝方无人机都能成功突防。很容易知道蓝方遭遇红方拦截的场景有两种，分别是遭遇一个无人机集群的拦截，和两个无人机集群的围堵，下面针对这两种情况分析蓝方无人机的躲避策略。当无人机前方只有一个无人机集群拦截时，蓝方躲避的飞行轨迹如下图 3 所示。

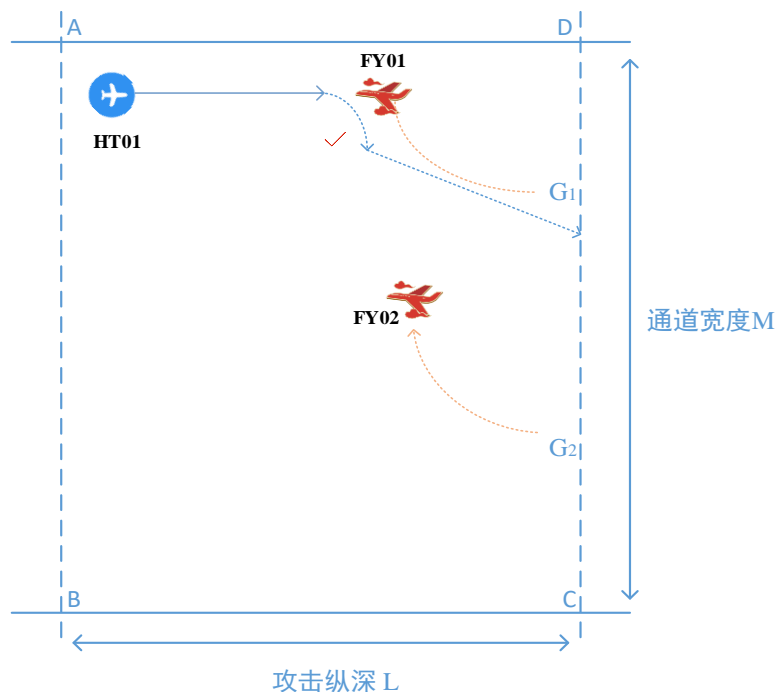


图 3 一个红方无人机机群拦截场景图

由于速度优势，蓝方无人机通过向侧边转弯一定能够躲避单个无人机集群的拦截。

当两个红方无人机集群同时出现在蓝方无人机前方进行夹击时，场景图如下图 4 所示，蓝方无人机无法逃脱。

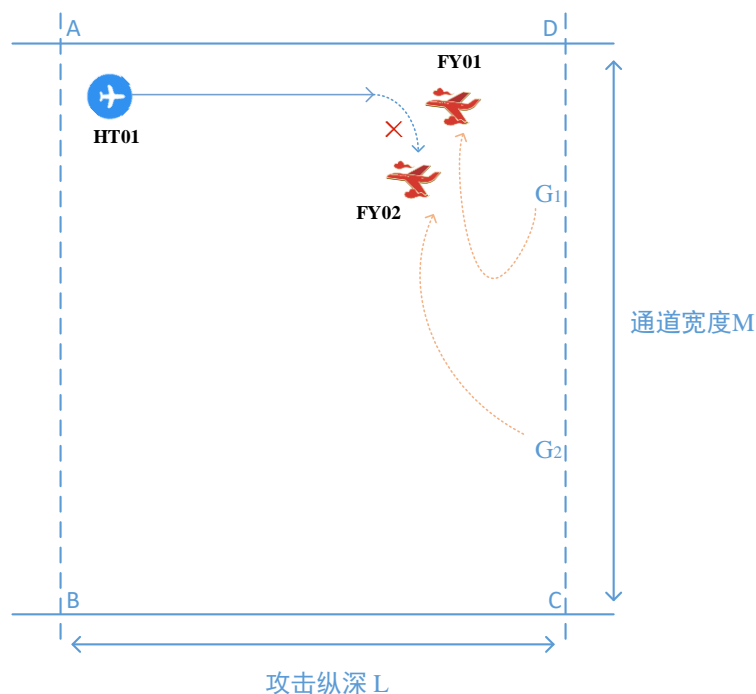


图 4 两个红方无人机机群夹击场景图

仍然假定蓝方无人机水平直线飞行，红方两个无人机集群可以实时探测到彼此的位置，从而“约定”好同时抵达蓝方无人机的前方。以上图 4 中的蓝方无人机 HT01 飞行轨迹线为例，无人机集群 FY01 一定比无人机集群 FY02 率先碰到蓝方无人机，它可以在合理的飞

行区域内绕行以拖延时间保证与无人机集群 FY02 汇合，如上图 4 所示。那么此时只要考虑无人机集群 FY02 是否能及时抵达蓝方无人机航线处进行拦截，如果无法及时抵达，红方两个无人机集群就不能同时出现在蓝方无人机的前方夹击，拦截失败。

本题中两个红方无人机集群的起始点位置均已给定，FY02 初始位置为 G_2 点 $(50 \times 10^3, 20 \times 10^3)$ ，假设蓝方无人机从上边界 AD 上某点处开始起飞，起点 E 的坐标为 $(x_{b0}, 70 \times 10^3)$ 处，对抗场景如下图所示 5 所示。

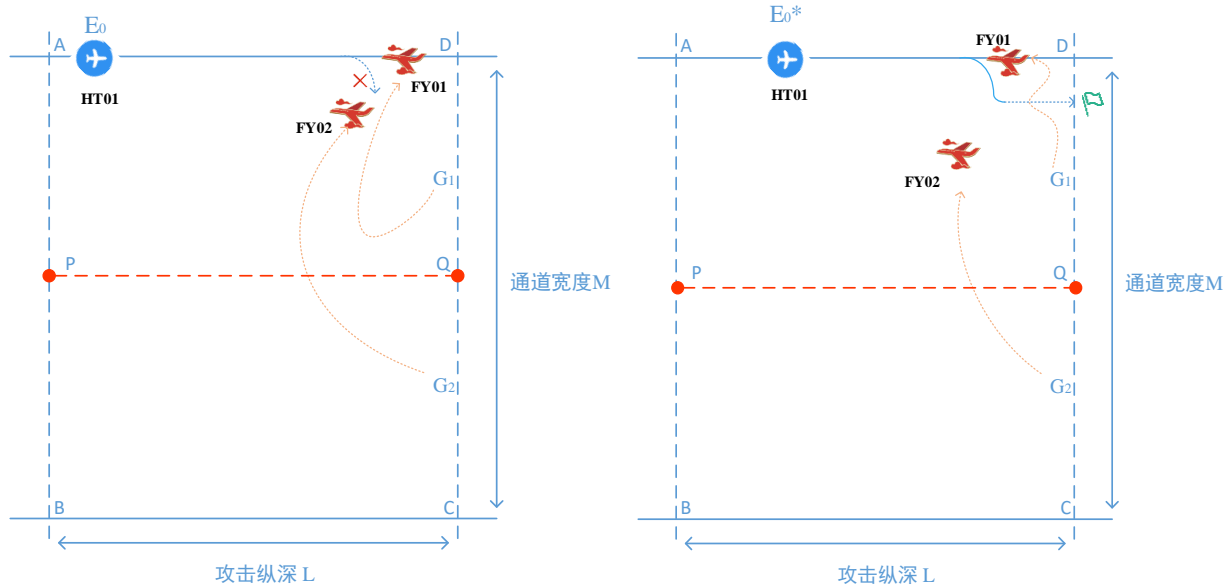


图 5 蓝方无人机突防场景图

在图 5 左侧场景图中，HT01 从 E_0 点出发，飞行一段时 t_0 后，红方两个无人机集群汇合夹击蓝方无人机，此时蓝方突防失败；在右侧场景图中，存在一个突防临界点 E_0^* 点，HT01 出发后不会遭遇红方夹击，即红方无人机集群 FY02 无法及时到达 HT01 航线附近与 FY01 汇合，对蓝方实施围堵，此时拦截失败。那么在线段 AD 上 E_0^* 右侧所有点处，蓝方无人机都能成功突防。因此本题模型求解的目标就是找到矩形区域 ABCD 内每一条水平线上的临界点 E_0^* ，利用它们拟合出一条临界线，在此临界线与右边界 DC 所包围区域内的所有位置上，无论红方无人机采取什么策略，蓝方无人机都能成功躲避拦截，完成土方任务。

5.1.3 模型求解

对于本题路径规划模型的求解，首先要依据前提条件（比如飞行速度，初始位置等），利用相邻两点的无人机位置坐标关系拟合出三个点的移动轨迹线方程。通过分析在同一时刻轨迹线是否相交来判断无人机是否能成功突防或者拦截，如果找不到这样的时间 t ，则认为红蓝双方无法相遇。

1、理论分析

根据红、蓝双方相邻两个时刻的位置坐标，令 $\lim \Delta t \rightarrow 0$ ，可以求得他们的飞行轨迹线方程 $l_{HT01}(x, y)$, $l_{FY01}(x, y)$, $l_{FY02}(x, y)$ ，分别由关系式 (1) (2) (3) 获得。要求临界点，只关心红方无人机集群 FY01 或 FY02 的轨迹方程 $l_{FY01}(x, y)$ 、 $l_{FY02}(x, y)$ 和蓝方无人

机 HT01 的轨迹方程 $l_{HT01}(x, y)$ 是否相遇。

在上半区域 APQD 中，只关心 $l_{FY02}(x, y)$ 和 $l_{HT01}(x, y)$ 是否在同一时刻相交，即判断 $l_{FY02}(x, y) = l_{HT01}(x, y)$ 中 t 是否有解，当根的判别式 $\Delta < 0$ 时无解。那么只要找到使得 $\Delta = 0$ 的蓝方无人机起点 E_0^* 坐标 $(x_{b0}, 70 \times 10^3)$ ，就找到了其无论如何都能成功突防得的临界点。下半区域 PBCQ 只关心轨迹线方程 $l_{FY01}(x, y)$ 和 $l_{HT01}(x, y)$ ，分析思路相同。容易知道，上下两半区域内的可突防域关于线段 PQ 对称的，因此只要求解出任意半个区域内的可突防域即可。我们下面分析矩形区域 APDQ。

2、预想结果

从已知红方无人机集群的圆周半径为 100m，为了方便求解，我们在竖直方向上，每隔 100m 取一条与边界 AD 平行等长的线段作为蓝方无人机的起飞路径，则在矩形区域 APDQ 内需要取 N 条轨迹线寻找 N 个临近点，其中 $N = (70 \times 10^3 m \div 2) \div 100m = 350$ 。利用这 350 个点可以拟合出若干条平滑的临界线，与右侧边界 CD 围成的区域内任意位置，都是蓝方的可突防域。

对区域离散化出 350 条水平线可以大大减少程序计算复杂度，按照理论推导的结果大致如下图 6 所示。

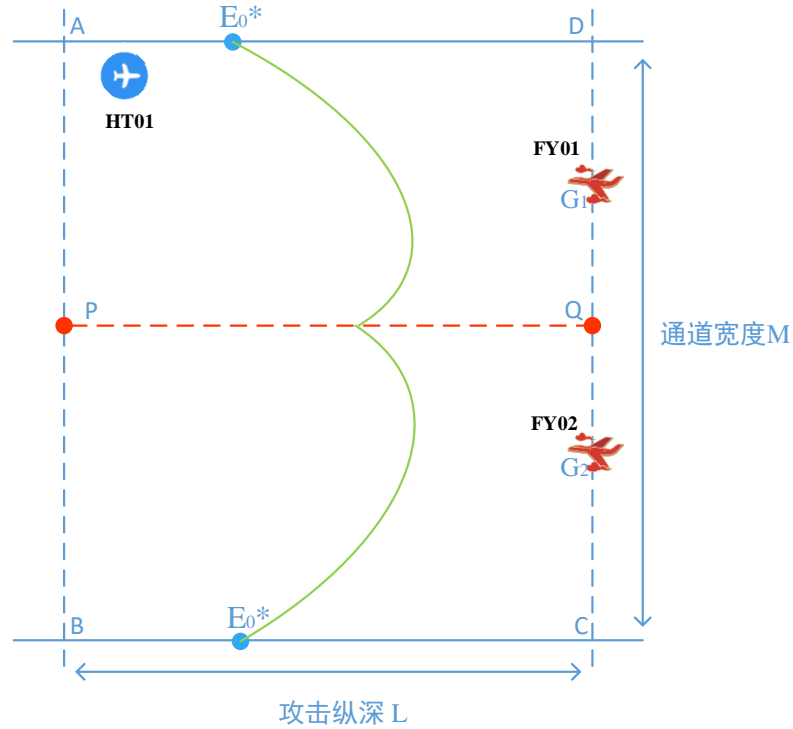


图 6 预想结果图

3、仿真结果

仿真时采用热力图的形式来表示红方无人机集群对蓝方无人机的威胁程度，不同位置上有不同大小的威胁值，因而得到的结果是平面区域内威胁程度分布情况。如下图 7 所示，仿真结果与预想结果基本吻合，在临界区域仿真结果出现不确定性，更能对应实际对抗场景。

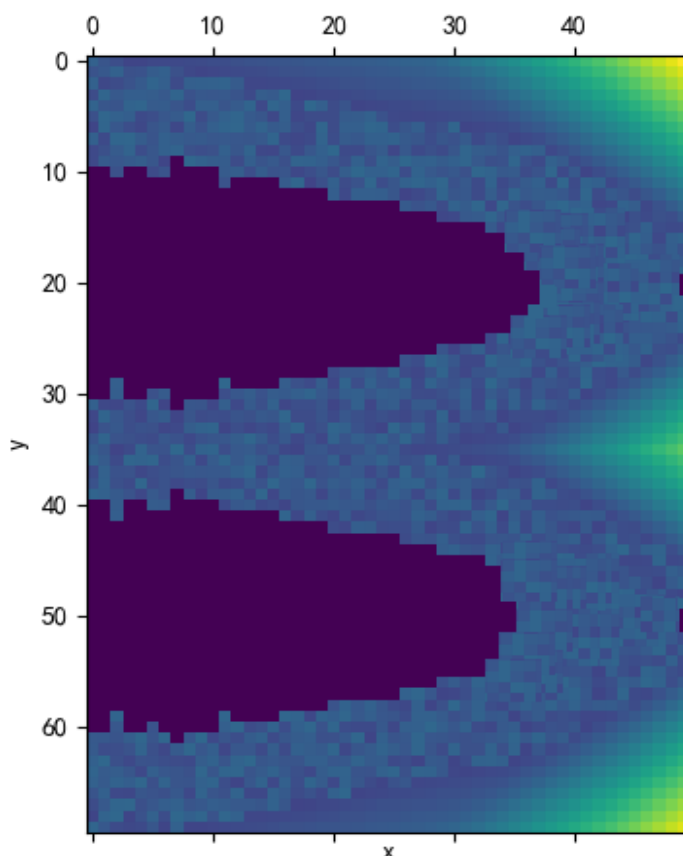


图 7 仿真结果图

对于图 7，蓝色区域表示蓝方无人机有被红方无人机集群拦截的危险，颜色越深威胁程度越高，即位置越向左侧靠近越有可能被拦截；黄色区域表示蓝方可以安全通过完成突防，颜色越亮越安全；中间蓝色到黄色的渐变区域表示蓝方能以一定概率突防成功，即当红方采取某些不当围堵策略时，蓝方可以突防成功。

4、蓝方最优突防策略

根据以上分析内容，可以总结出蓝方在可突防域内的最优突防策略，简单示意图如下图 5 右侧图所示。

当前方没有威胁或者威胁较小时，蓝方无人机应当尽量保持水平直线飞行，才能保证以最短的时间越过边界 CD，完成突防；当中途遇到红方无人机拦截时，蓝方应当在敌方靠近时迅速转弯躲避，转过的弧线刚好与红方无人机集群的圆周相切，并且在成功躲避红方后再次转弯至水平方向直行飞越右侧边界 CD。此时蓝方无人机在红方任意拦截策略下都能保证以最短的时间完成突防任务。我们将所得的最佳策略总结如下：

- 1) 强行突破：以 CD 边界上距离最近的点为目标。
- 2) 边路策略：以距离最近的对角（C 或 D）为目标。
- 3) 欺骗策略：先执行边路策略，待红机接近，反向突破后执行强突策略。
- 4) 根据战场态势综合应用。

最后，根据仿真结果和突防策略我们将突防策略与战场位置做如下图 8 映射。

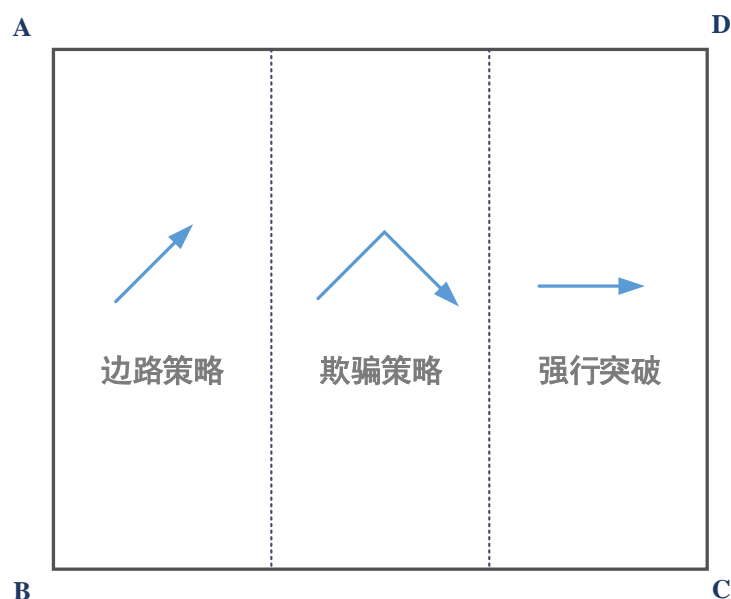


图 8 蓝方突防策略图

5.2 问题二：微分博弈模型

5.2.1 模型建立

双方无人机在飞行过程中，均受到最小转弯半径的控制约束，另外，蓝方无人机还受到边界约束。我们假设：1) 在整个过程在蓝方无人机被拦截时结束；2) 红方两个无人机集群的中心相距 12km，且 $G1$ 和 $G2$ 的垂直平分线过蓝方无人机的初始位置。

我们通过微分博弈对红蓝双方的控制变量 $u(t), v(t)$ 进行博弈，通过构造哈密顿函数建模，寻找模型求解的鞍点，从而得到最优决策 $u^*(t), v^*(t)$ 。在本题目中，控制变量设置为无人机偏转角度， $u(t) = \theta_E(t), v(t) = \theta_P(t)$ 。原理图如下图 8 所示。

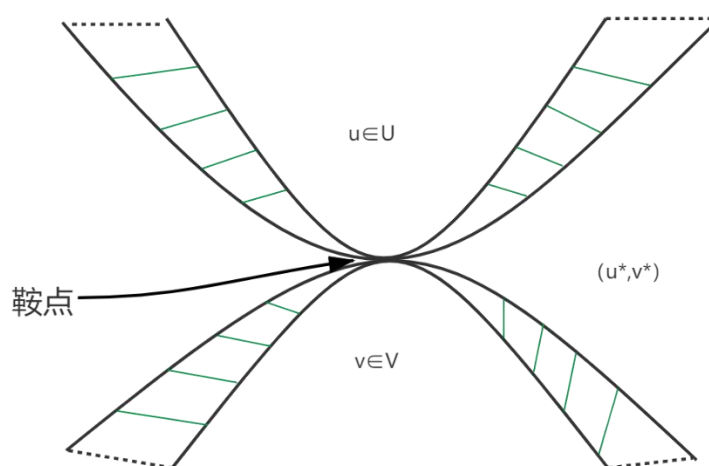


图 9 微分博弈模型原理图

1、确定优化目标函数及约束条件

在本题目中，最小化蓝方无人机从出发到被红方拦截位置的垂直距离，是我们的优化

目标。即在控制和边界的约束下，我们希望得到蓝方无人机从出发位置到被红方拦截位置之间的最短垂直高度。

目标函数为：

$$J = M + \beta * t_M \quad (4)$$

其中， M 为通道带宽， t_M 为蓝方无人机到达边界 CD 的时间， β 为 M 与 t_M 的相关系数，用于调整 M 与 t_M 之间的大小。求解模型，我们首先要得到的是通道带宽的下限 M_{min} ，在此基础上再去求最短的突防时间 t_{Mmin} ，因此设置相关系数时，要使 $M \gg \beta * t_M$ 。我们设置 $\beta = 0.01$ 。

蓝方无人机被拦截的条件：

$$\Psi = (y_E(t_M) - \frac{1}{2}M)^2 + (y_{P1}(t_M) - \frac{1}{2}M)^2 + (y_{P2}(t_M) - \frac{1}{2}M)^2 = 0 \quad (5)$$

其中， $y_E(t_M)$ 表示蓝方无人机在 t_M 时刻的纵向高度， $y_{P1}(t_M)$ 表示红方一个无人机集群在 t_M 时刻的纵向高度， $y_{P2}(t_M)$ 表示红方另一个无人机集群在 t_M 时刻的纵向高度。

终端时刻约束：

$$\Phi = \phi + v^T * \Psi = y_E(t_M) + a * \Psi \quad (6)$$

其中， ϕ 为终端指标， a 为拉格朗日乘子。

边界约束条件：

$$y_E(t) = 0 \text{ or } M \quad (7)$$

控制约束条件：

$$\begin{cases} R_E > r_{Emin} = 500 \\ R_P > r_{Pmin} = 350 \end{cases} \quad (8)$$

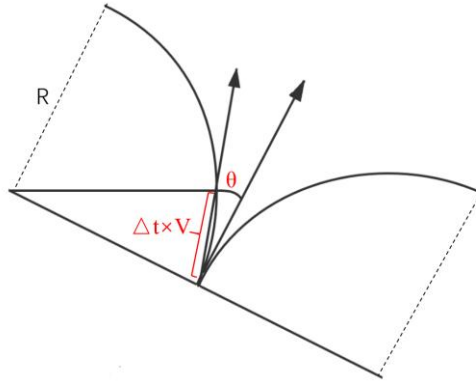


图 10 偏转角与圆周半径关系图

我们设置时间间隔为 $t=1$ 秒，计算在 1 秒内蓝方与红方无人机的速度偏转角度 θ_E 和 θ_P ，并假设 $t=1s$ 内蓝方或红方无人机沿着 θ_E 或 θ_P 方向直线飞行，飞行距离为 $R'_E = V_E * t = 250m$ 、 $R'_P = V_P * t = 200m$ 。由几何关系得知，我们假设最小和最大偏转角 θ_{Emin} 和 θ_{Pmin} 都等于 0° ，则计算得 θ_{Emin} 和 θ_{Pmin} 都约等于 $+36^\circ$ 。故以上控制约束条件又可以转化为：

$$\begin{cases} \theta_{Emin} < \theta_E = \arcsin \frac{R_E}{R'_E/2} < \theta_{Emax} \\ \theta_{Pmin} < \theta_P = \arcsin \frac{R_P}{R'_P/2} < \theta_{Pmax} \end{cases} \quad (9)$$

5.2.2 模型求解

1、构建哈密顿函数

$$H = \lambda_1 * V_E * \cos\theta_E + \lambda_2 * V_E * \sin\theta_E + \lambda_3 * V_P * \cos\theta_P + \lambda_4 * V_P * \sin\theta_P + \mu_1 * (R_E - R_{Emin}) + \mu_2 * (R_P - R_{Pmin})$$

其中, $\lambda_i (i = 1, 2, 3, 4)$ 为协态变量, 且拉格朗日因子 $\mu_i (i = 1, 2)$ 定义如下:

$$\mu_1 \begin{cases} = 0, & R_E \leq R_{Emin} \\ \neq 0, & R_E = R_{Emin} \end{cases} \quad \mu_2 \begin{cases} = 0, & R_P \leq R_{Pmin} \\ \neq 0, & R_P = R_{Pmin} \end{cases} \quad (10)$$

2、求解最优策略

由哈密顿函数对各状态 $X = [y_E, V_E, y_P, V_P]^T$ 求偏导得到协态方程组:

$$-\frac{\partial H}{\partial x} = \begin{cases} 0 \\ \lambda_1 * \cos\theta_E + \lambda_2 * \sin\theta_E \\ 0 \\ \lambda_3 * \cos\theta_P + \lambda_4 * \sin\theta_P \end{cases} \quad (11)$$

根据微分博弈中的双边极值定理, 令哈密顿函数对控制变量 θ_E 和 θ_P 求一阶偏导数:

$$\frac{\partial H}{\partial \theta_E} = -\lambda_1 * V_E * \sin\theta_E + \lambda_2 * V_E * \cos\theta_E + \mu_1 * (R'_E/2) * \cos\theta_E \quad (12)$$

$$\frac{\partial H}{\partial \theta_P} = -\lambda_3 * V_P * \sin\theta_P + \lambda_4 * V_P * \cos\theta_P + \mu_2 * (R'_P/2) * \cos\theta_P \quad (13)$$

我们定义

$$S_1 = -\lambda_1 * V_E * \sin\theta_E + \lambda_2 * V_E * \cos\theta_E = \mu_1 * (R'_E/2) * \cos\theta_E \quad (14)$$

$$S_2 = -\lambda_3 * V_P * \sin\theta_P + \lambda_4 * V_P * \cos\theta_P = \mu_2 * (R'_P/2) * \cos\theta_P \quad (15)$$

综上可知

$$\begin{cases} \theta_E = \theta_{Emin}, & S_1 < 0 \\ \theta_E = \theta_{Emax}, & S_1 > 0 \\ \theta_{Emin} < \theta_E < \theta_{Emax}, & S_1 = 0 \end{cases} \quad (16)$$

$$\begin{cases} \theta_P = \theta_{Pmin}, & S_2 < 0 \\ \theta_P = \theta_{Pmax}, & S_2 > 0 \\ \theta_{Pmin} < \theta_P < \theta_{Pmax}, & S_2 = 0 \end{cases} \quad (17)$$

其中, $\theta_{Emin} = \theta_{Pmin} = 0$, $\theta_{Emax} = \theta_{Pmax} = 36^\circ$ 。

令哈密顿函数对控制变量求二阶导数得:

$$\frac{\partial^2 H}{\partial \theta_E^2} = -\lambda_1 * V_E * \cos\theta_E - \lambda_2 * V_E * \sin\theta_E - \mu_1 * (R'_E/2) * \sin\theta_E \quad (18)$$

$$\frac{\partial^2 H}{\partial \theta_P^2} = -\lambda_3 * V_P * \cos\theta_P - \lambda_4 * V_P * \sin\theta_P - \mu_2 * (R'_P/2) * \sin\theta_P \quad (19)$$

由二阶最优条件: $\frac{\partial^2 H}{\partial u^2} \geq 0$, $\frac{\partial^2 H}{\partial v^2} \leq 0$, 其中 $u = \theta_E$, $v = \theta_P$ 。根据红方期望哈密顿函数值最小, 蓝方希望哈密顿函数值最大, 综合一阶最优条件和二阶最优条件可以得到红蓝双方的最优控制策略为:

$$\theta_E^* = \begin{cases} \theta'_E, & S_1 = 0 \\ \theta_{Emin}, & S_1 < 0 \\ \theta_{Emax}, & S_1 > 0 \end{cases} \quad (20)$$

$$\theta_P^* = \begin{cases} \theta'_P, & S_2 = 0 \\ \theta_{Pmin}, & S_2 < 0 \\ \theta_{Pmax}, & S_2 > 0 \end{cases} \quad (21)$$

其中, θ'_E 、 θ'_P 为 $S_1 = 0$ 、 $S_2 = 0$ 对 θ_E 、 θ_P 的解。

综上, 我们可以将原问题构建成为一个两点边值问题, 通过求解此两点边值问题即可证明并得到通道带宽 M 的下限 M_{min} , 当实际通道带宽比 M_{min} 大时, 蓝方无人机一定能突破红方无人机集群的拦截, 同时得到给出此种情形下蓝方无人机时间最短的突防策略。

5.2.3 实验结果分析

我们通过微分博弈的方法, 在蓝方无人机在目标通道宽度 M 尽可能小的情况下, 躲避红方无人机集群的拦截。首先通过构造两点边值问题, 然后求解问题寻找鞍点及鞍点对应的最佳策略。通过对模型的求解, 我们得到了蓝方无人机的最优突防策略, 在此策略下, 蓝方无人机刚好能躲过红方无人机群的拦截, 成功突防。

我们通过实验证明存在一个通道带宽的下限 M_{min} , 当实际通道带宽大于 M_{min} 时, 蓝方无人机一定能成功突防。我们记录了蓝方无人机在整个过程中飞行轨迹的最大纵向飞行距离为 16.1km。因为蓝方无人机从 AB 边界中点出发, 所以通道带宽最小为 $M_{min} = 2 * 16.1km = 32.2km$ 。

蓝方无人机从边界 AB 的中点出发, 根据求解模型得到的时间最短突防策略以及红方最佳防守策略:

$$\theta_E^* = \begin{cases} \theta'_E, & S_1 = 0 \\ \theta_{Emin}, & S_1 < 0 \\ \theta_{Emin}, & S_1 > 0 \end{cases} \quad (22)$$

$$\theta_P^* = \begin{cases} \theta'_P, & S_2 = 0 \\ \theta_{Pmin}, & S_2 < 0 \\ \theta_{Pmin}, & S_2 > 0 \end{cases} \quad (23)$$

记录蓝方无人机以及红方无人机集群的飞行轨迹, 如图 11 所示。我们从实验仿真结果可以看出, 蓝方无人机刚开始一直进行直线飞行, 吸引红方无人机集群靠拢, 直到它们足够接近, 蓝方无人机开始转弯, 利用其自身的速度优势, 成功躲开了红方无人机的包夹拦截。

如果在同一时刻, 蓝方无人机所在位置的横坐标大于红方的任意一个无人机集群的横坐标, 利用自身速度优势, 蓝方无人机即可成功突防。因为蓝方无人机的速度比红方无人机的速度大, 所以当蓝方无人机转弯成功绕过红方无人机群之后, 红方无人机不可能再次对蓝方无人机进行包夹拦截。

在我们的轨迹记录实验结果中, 蓝方无人机刚开始是进行斜向直线飞行, 这是为了尽快缩短红方两个集群相互靠近的时间。随后出现了三次拐弯的情况, 我们分析: 第一次拐弯是为了尽快接近红方无人机群; 第二次拐弯是因为第一个红方无人机集群与蓝方无人机的距离足够接近时, 蓝方无人机预判出继续直线前行必然会与红方无人机相遇, 因此蓝方做出了相应调整以达到绕过红方无人机群的效果; 第三次拐弯也是为了避免不被红方的第二个无人机群拦截。当蓝方无人机位置的横坐标已经超过红方的两架无人机后, 红方必定无法对蓝方无人机进行拦截, 蓝方无人机可成功突防。

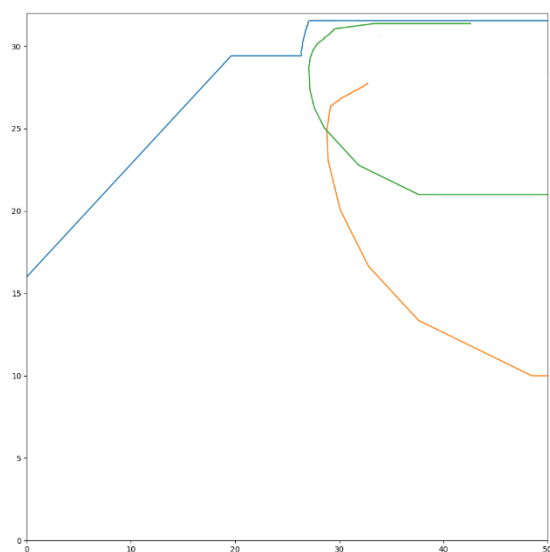


图 11 蓝方无人机成功突防：蓝线表示蓝方无人机轨迹，橘线和绿线表示红方无人机集群轨迹

我们进一步对模型求解结果进行验证，当实际通道带宽比 M_{min} 小时，会出现蓝方无人机被拦截的情况。设置实际通道带宽为 $M = M_{min} - 2km$ 。我们再次记录各无人机的飞行轨迹，发现蓝方无人机被拦截。从而证明了通道带宽小于 $M_{min} = 32.2km$ 时，存在策略使得红方无人机集群能够成功拦截蓝方无人机，即证明了 M_{min} 是蓝方无人机一定能突破红方无人机集群的下限。

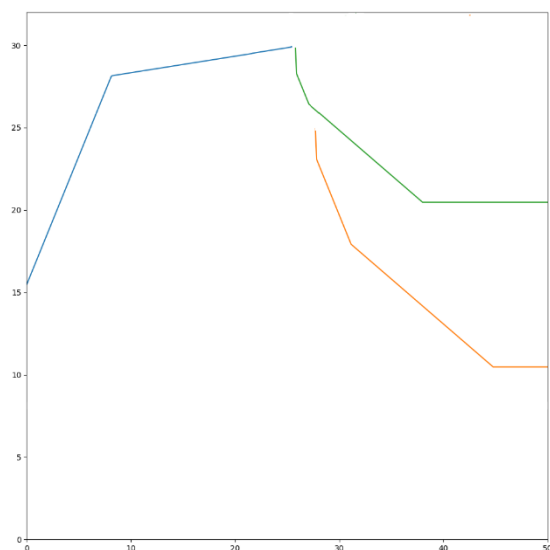


图 12 蓝方无人机被成功拦截：蓝线表示蓝方无人机轨迹，橘线和绿线表示红方无人机集群轨迹

5.3 问题三：改进的微分博弈模型

5.3.1 模型建立

我们假定红方的两个运载机相距 30km。红方的运载机可分两个波次发射 10 架无人机，我们假定红方的每次波次都发送 5 架无人机。

我们以问题一中得到的划分可拦截区域与不可拦截区域的边界，当蓝方无人机到达安

全区域即不可拦截区域，大概距离红方无人机集群 15km 处时，红方运载机开始发射第二个无人机集群。

1、确定优化目标函数及约束条件

本题目中，蓝方无人机需要躲避无人机集群数量为 4 个，优化目标是寻找一个一个通道带宽 M 的下限。因此，设置目标函数为：

$$J = M \quad (24)$$

蓝方无人机被拦截的条件：

$$\begin{aligned} \Psi = & (y_E(t_M) - \frac{1}{2}M)^2 + (y_{P11}(t_M) - \frac{1}{2}M)^2 + (y_{P12}(t_M - t_f) - \frac{1}{2}M)^2 + (y_{P21}(t_M) - \\ & \frac{1}{2}M)^2 + (y_{P22}(t_M - t_f) - \frac{1}{2}M)^2 = 0 \end{aligned} \quad (25)$$

其中， $y_E(t_M)$ 表示蓝方无人机在 t_M 时刻的纵向高度， $y_{P11}(t_M - t_f)$ 表示红方一个无人机集群第一波次在 t_M 时刻的纵向高度， $y_{P2}(t_M)$ 表示红方另一个无人机集群在 t_M 时刻的纵向高度， t_f 为蓝方无人机进入第二问中的不可拦截区域时。

终端时刻约束：

$$\Phi = \phi + v^T * \Psi = y_E(t_M) + a * \Psi \quad (26)$$

其中， ϕ 为终端指标， a 为拉格朗日乘子。

边界约束条件：

$$y_E(t) = 0 \text{ or } M \quad (27)$$

控制约束条件：

$$\begin{cases} \theta_{Emin} < \theta_E = \arcsin \frac{R_E}{R'_E/2} < \theta_{Emax} \\ \theta_{Pmin} < \theta_P = \arcsin \frac{R_P}{R'_P/2} < \theta_{Pmax} \end{cases} \quad (28)$$

其中， θ_E 和 θ_P 为蓝方与红方无人机转弯时的偏转角度， R_E 和 R_P 为蓝方和红方无人机转弯时的圆周半径。

5.3.2 模型求解

模型求解过程与问题二基本相同，得到红蓝双方的最优控制策略为：

$$\theta_E^* = \begin{cases} \theta'_E, & -\lambda_1 * V_E * \sin\theta_E + \lambda_2 * V_E * \cos\theta_E = \mu_1 * (R'_E/2) * \cos\theta_E \\ \theta_{Emin}, & -\lambda_1 * V_E * \sin\theta_E + \lambda_2 * V_E * \cos\theta_E < \mu_1 * (R'_E/2) * \cos\theta_E \\ \theta_{Emin}, & -\lambda_1 * V_E * \sin\theta_E + \lambda_2 * V_E * \cos\theta_E > \mu_1 * (R'_E/2) * \cos\theta_E \end{cases} \quad (29)$$

$$\theta_P^* = \begin{cases} \theta'_P, & -\lambda_3 * V_P * \sin\theta_P + \lambda_4 * V_P * \cos\theta_P = \mu_2 * (R'_P/2) * \cos\theta_P \\ \theta_{Pmin}, & -\lambda_3 * V_P * \sin\theta_P + \lambda_4 * V_P * \cos\theta_P = \mu_2 * (R'_P/2) * \cos\theta_P \\ \theta_{Pmin}, & -\lambda_3 * V_P * \sin\theta_P + \lambda_4 * V_P * \cos\theta_P = \mu_2 * (R'_P/2) * \cos\theta_P \end{cases} \quad (30)$$

其中， θ'_E 、 θ'_P 为 $-\lambda_1 * V_E * \sin\theta_E + \lambda_2 * V_E * \cos\theta_E = 0$ 、 $-\lambda_1 * V_E * \sin\theta_E + \lambda_2 * V_E * \cos\theta_E = 0$ 对 θ_E 、 θ_P 的解。

综上，我们构建一个两点边值问题，通过求解此两点边值问题即可得到通道带宽 M 的上限 M_{max} ，当实际通道带宽比 M_{max} 小时，无论蓝方无人机采用什么样的突防策略，对于红方无人机集群均存在相应的拦截策略，在尽量小的通道带宽内成功拦截蓝方无人机。

5.3.3 实验结果分析

我们通过实验证明，存在着一个通道带宽的上限 M_{max} ，使得当实际带宽小于 M_{max} 时，红方无人机集群总能对蓝方无人机进行拦截。我们记录了蓝方无人机在整个过程中飞行轨迹的最大纵向飞行距离为 $\frac{1}{2} * M_{max} = 34.2km$ 。因为蓝方无人机从 AB 边界中点出发，所以通道带宽最小为 $M_{max} = 2 * 34.2km = 68.4km$ 。

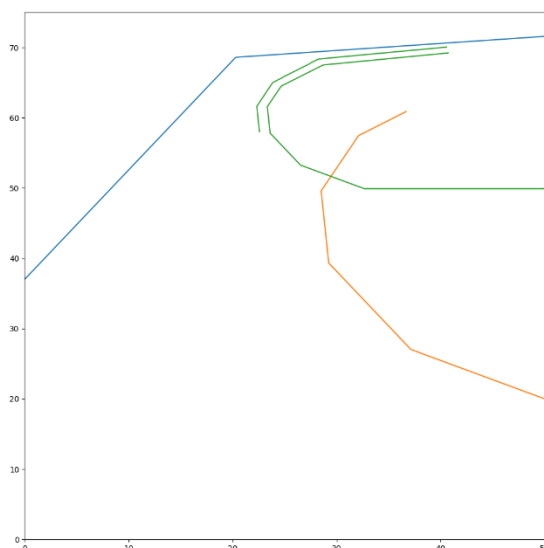


图 12 蓝方无人机成功突防：蓝线表示蓝方无人机轨迹，绿线和橘线分别表示红方两个运载机发射的无人机集群

当蓝方无人机在纵向方向飞行了 $0.5 * M_{max}$ 时，刚好被红方无人机集群拦截，记录此过程中蓝方无人机和红方各个无人机集群的飞行轨迹，如图 12 所示。我们由图中可以看出，当蓝方无人机开始进入问题一中的不可拦截区域时，离蓝方无人机较近的红方运载机释放出第二波次的无人机集群，并与其初始发射的无人机集群共同对蓝方无人机进行拦截，并成功阻止了蓝方无人机的突防。

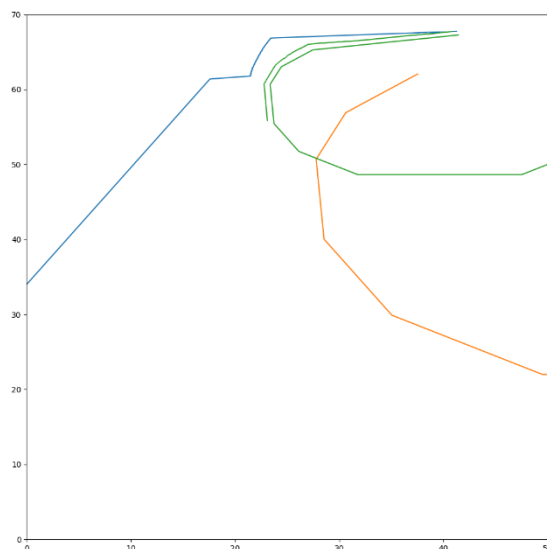


图 13 红方无人机集群成功拦截：蓝线表示蓝方无人机轨迹，绿线和橘线分别表示红方两个运载

机发射的无人机集群

如上图 13 所示,通过实验再次证明,如果实际通道带宽小于 M_{max} ,红方无人机集群均存在相应的拦截策略,能够成功阻止蓝方无人机的突防。我们设置 $M = M_{max} - 5km$,记录无人机的飞行轨迹。实验证明,当实际通道带宽小于 M_{max} 时,红方无人机集群均能实现对蓝方无人机的集群。

5.4 基于 BP 神经网络的目标分配及博弈模型

5.4.1 方法分析

我们将问题四分析成一个多机协同目标分配的博弈对抗问题。蓝方有三架无人机,红方有五架运载机,分别能产生两个无人机集群,即红方一共有十个无人机集群。问题的关键点在于十个无人机集群怎样分配,同时拦截下蓝方三架无人机;拦截效果还与分配后的协同策略以及蓝方的协同突防策略息息相关,因此这这也是一个博弈问题,并且本题背景也存在较多的不确定性。

1、多机目标分配问题

首先,多无人机协同任务规划即是根据一组特定条件的约束,以实现预定准则函数的最优或次优为目标,将作战任务分解成一些子任务并分配给各个无人机,主要关心任务分配和路径规划。任务分配考虑各种约束条件,以总任务有效达成为目标,将具体目标和行动任务分配给各机,而各机根据分配的任务再进行具体的作战路径规划;路径规划的功能是在满足如最大飞行速度、最小转弯半径、完成任务最短时间等自身或外部限制的前提下在一系列位置之间设计或生成路径。现实中类似的问题也有很多,这种以集中式控制系统为基础进行任务分配最常用的模型有多旅行商问题 MTSP 模型^[6]、车辆路径问题 VRP 模型^[7]、多机路径分配 mVRP 模型^[8]、混合整数线性规划问题 MILP 模型^[9]、动态网络流优化 CMTMP 模型^[10]、基于演化深度神经网络的无人机协同无源定位动态航迹规划模型^[11]等。传统模型中无人机集群需要一个中心站用于目标位置解算以及给出下一时刻各个无人机航迹点,属于集中式控制,随着集群内无人机数目的增多,集群网络的鲁棒性降低。因此本题借鉴参考文献^[10]中的思想,初步选择较为基础的 BP 神经网络训练红、蓝双方无人机,从而完成多机协同目标分配任务。

2、双方协同博弈问题

在红、蓝方无人机进行对抗时,想要双方都采取最优策略需要在博弈均衡的情况下实现。实际对抗场景下红方和蓝方都没有绝对的最优策略,两者需要进行不断的对抗、学习和改进,从而不断优化自己的作战策略,直到达到一种均衡的结果。类似的博弈对抗模型也有很多,比如谷歌历时数年针对星际争霸游戏开发的 AlphaStar 项目,就是一个大型的人工智能系统,其中借助多智能体的思想,使用多种神经网络训练不同类型的作战智能体并进行级联。但是星际争霸游戏的战场环境相对于本题要复杂很多,借鉴意义较小;值得参考的还有今年 8 月 20 日刚刚结束的美国 DARPA AlphaDogfight 挑战赛,赛题背景也是基于红、蓝双方的战斗机对抗问题。因此本题在求解时使用强化学习思想,训练红、蓝双方无人机,从而做出相对最优的拦截策略和突防策略。

5.4.2 基于 BP 神经网络的目标分配模型

1、模型建立

BP 神经网络具有很强的非线性映射能力，不需要任何先验公式，当隐层的神经元节点规模足够大时，它就可以在任意误差范围内逼近最佳系统的非线性函数。而本题中很多先验信息都未给定，比如红蓝双方无人机的起飞位置和飞行策略、红方第二波次无人机的起飞时刻等都未定，因此很难使用传统的目标分配模型进行求解，这种情况下使用神经网络显得更为合理。BP 神经网络的基本模型如下图 13 所示。

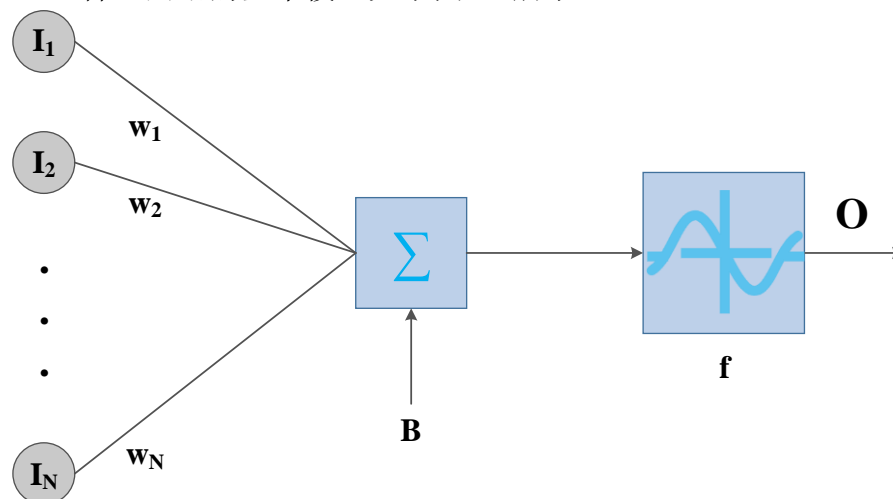


图 14 BP 神经网络的基本模型

上图中， $I_1 \sim I_N$ 为 N 个输入变量， $w_1 \sim w_N$ 为输入神经元的权值， B 为神经元的阈值， f 为网络的激活函数， O 为神经网络的输出。此为三层神经网络，包含输入层、隐含层、输出层，更多层 BP 神经网络通常构建为级联结构，所有中间网络层统称为隐藏层 H 。

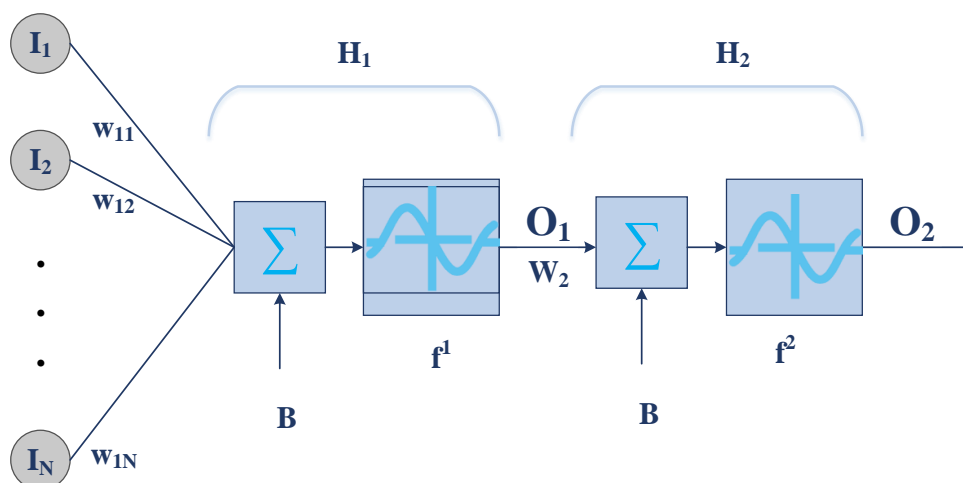


图 15 BP 神经网络的基本模型

上图 15 是一个四层神经网络，可见其实两个隐藏层 H_1 、 H_2 的级联。一般隐藏层数越多，非线性拟合能力越强，比如经常被使用的深度神经网络，但是随着层数和节点数的增加，网络的计算复杂度也大大提升。对于简单问题，可以使用相对轻量级的网络架构，处

理性能更好。因此本题初步采用三层 BP 网络结构学习训练多机协同目标分配任务，解决如何针对蓝方三架无人机分配兵力的问题，模型图如下图 16 所示。

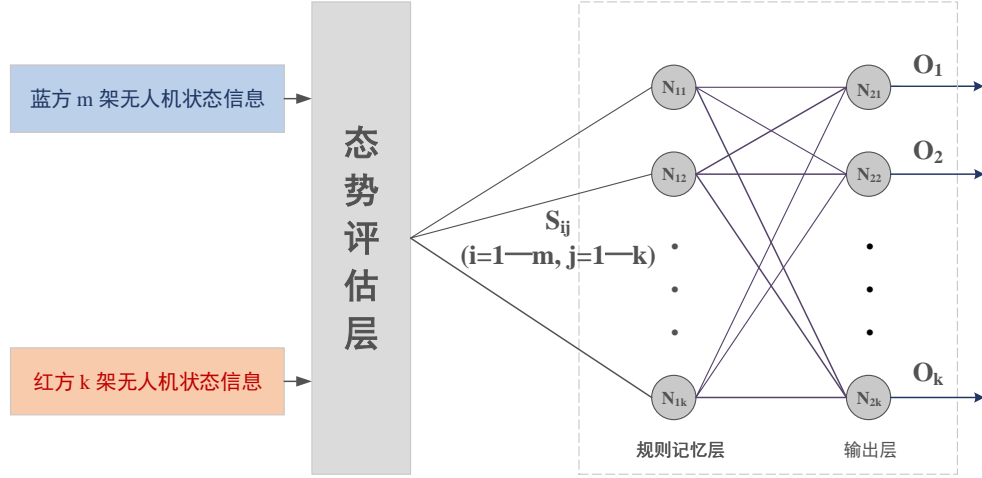


图 16 协同目标分配的 BP 神经网络模型

对于本题而言， $m=3, k=5 \times 10=50$ ，那么模型建立具体思想为：

1) 输入层：将蓝方 3 架无人机和红方 50 架无人机的状态信息 $[f, h, v, T_{dc}]$ 输入到态势评估层，根据自主优先权的公式计算攻击优势函数 S_{ij} ，其中 f 、 h 为无人机的位置信息， v 为速度信息， T_{dc} 为拦截能力指数。那么态势评估层输出的 S_{ij} 即为神经网络的输入。

要确定符合本题设的拦截优势函数，首先要确定其影响因素，所以我们定义了距离优势函数和拦截效能优势函数。

① 距离优势函数

根据问题一的仿真结果发现在临界线附近存在一些不确定性区域，这可能与红方无人机采取拦截策略不同或者蓝方无人机的突防策略不同有关。为了简化问题方便建模，我们假设不确定性是由攻击距离造成的，不同的攻击距离会使得红方无人机在蓝方无人机附近以不同的概率拦截成功。

假设无人机的攻击有效范围满足： $D_{\min} \leq D \leq D_{\max}$ ，其中 D_{\min} 和 D_{\max} 是无人机有效攻击范围的最小值与最大值，这两个距离也可以根据图 7 中的渐变色不确定区域宽度得到。在红蓝双方无人机的相对距离 $D \leq D_{\max}$ 时距离优势较小；从较远处接近蓝方时，攻击

优势随着 D 的减小而增大，在 $D_0 = \frac{D_{\min} + D_{\max}}{2}$ 时攻击优势达到最大，此后继续靠近蓝方距离优势又会逐渐衰减。所以距离优势函数会表现为类高斯分布，定义函数式为

$$T_D = \exp\left[-\frac{D - D_0}{\sigma}\right] \quad (31)$$

式中，如果令 $T_D(R_{\max}) = T_D(R_{\min}) = 0.95$ ，则有 $\sigma = 2(D_{\max} - D_{\min})$ 。

② 拦截效能优势函数

无人机集群拦截能力指数 C 主要使用集群数量 N_1 、各集群中无人机数量 N_2 来衡量。计算方法为：

$$C = \ln N_1 + \ln N_2 \quad (32)$$

为确保 C 值的精度在 $[0, 1]$ 之间，需要对其作进一步处理以便计算红方无人机集群对蓝方无人机的拦截效能函数 T_C 。定义 C_F 为红方无人机集群的拦截能力指数， C_H 为蓝方无人机的突防能力指数，则效能函数可以构造为：

$$T_C = \begin{cases} 0, & C_F / C_H < 0.3 \\ 0.25, & 0.3 \leq C_F / C_H < 1 \\ 0.5, & C_F / C_H = 1 \\ 0.75, & 1 < C_F / C_H < 1.5 \\ 1, & C_F / C_H \geq 1.5 \end{cases} \quad (33)$$

所以红方无人机集群对蓝方无人机的拦截优势函数为

$$T_{DC} = k_D T_D + k_C T_C \quad (34)$$

式中， k_D 、 k_C 分别为距离优势函数 T_D 和效能优势函数 T_C 的加权系数，且 $k_D + k_C = 1$ 。这两个系数的具体值可由多次网路训练仿真实验或者相关专家评议确定。

2) 隐藏层（规则记忆层）：神经元个数为：

$$Num = \sqrt{L + M} + A \quad (35)$$

式中 L 为隐藏层输入单元数 $2k$ ， M 为输出神经元数 k ， A 为 1-10 之间的调节常数，其具体值根据训练样本个数和精度要求来确定。

隐藏层的输出为

$$O_{il} = f_1 \left(\sum_{i=1}^{2k} w_{li} \times S_{ij} + N_{il} \right) \quad l=1,2,\dots,Num \quad (36)$$

式中 O_{il} 为第 l 个神经元的输出， w_{li} 为隐藏层第 l 个神经元对第 i 个输入的权值， N_{il} 为隐藏层第 l 个神经元的输出阈值。

3) 输出层（目标分配层）：为提高拦截效能，每个无人机集群中的每架无人机都只能分配给蓝方的一架无人机，因此输出结果形式为红方每架无人机的所属集群以及其被指派去拦截的某架蓝方无人机。

$$\bar{O}_{2j} = f_2 \left(\sum_{i=1}^{NN} w_{ji} \times O_{il} + N_{2j} \right) \quad j=1,2,\dots,k \quad (37)$$

$$O_{2j} = \text{floor}(\bar{O}_{2j}) \quad (38)$$

2、模型求解

根据题设，蓝方有三架无人机，红方每波次都派送五个集群，每个集群无人机数量不小于 3 架；蓝方无人机初始位置坐标是 $(0, y_{h-i})$ ， $i=1\sim 3$ ，红方无人机集群初始位置坐标为

$(0, y_{f-j})$, $j=1\sim 5$; 蓝方无人机飞行速度为 350m/s, 红方无人机飞行速度为 250m/s。红方无人机集群组合种类数为 $C_5^3=10$ 种, 每个集群中的无人机 两波次分配情况种类数为 5 种,

3-7、4-6、5-5、6-4、7-3。训练样本依据可能出现的对抗情形数制定:

- ① 按不同距离优势分别取 10 类: 不确定区域宽度均匀离散化出 10 个小距离值。
- ② 按不同效能优势分别取 50 类: 集群分配与两波次无人机架数分配可能情况数。

所以训练样本取以上两种优势类型的组合。激活函数 f^1 和 f^2 选用 BP 神经网络常用的 S 型函数, 反向传递过程中 使用的误差函数选用均方差函数, 再使用梯度下降法调整隐藏层中的权值和阈值。

依据上述设定进行仿真实验, 借助计算机训练神经网络, 学习红方兵力的最佳分配方式。

3、仿真结果

设置网络训练次数为 1000 次, 得到的目标分配结果如下表 2 所示:

表 2 目标分配表

蓝方无人机	红方集群	第一波次无人机数	第二波次无人机数
HT01	FY01	3	7
HT02	FY02-FY03-FY04	3-5-3	7-5-7
HT03	FY05	3	7

如表所示, 模型求解结果为: 无人机集群 FY01 拦截蓝方无人机 HT01, 第一波次发射 3 架无人机, 第二波次发射 7 架无人机; 无人机集群 FY02-FY03-FY04 拦截蓝方无人机 HT02, 第一波次分别发射 3、5、3 架无人机, 第二波次分别发射 7、5、7 架无人机; 无人机集群 FY05 拦截蓝方无人机 HT03, 第一波次发射 3 架无人机, 第二波次发射 7 架无人机。这种分配方案下, 拦截效果最好。作战场景图如下图 17 所示。

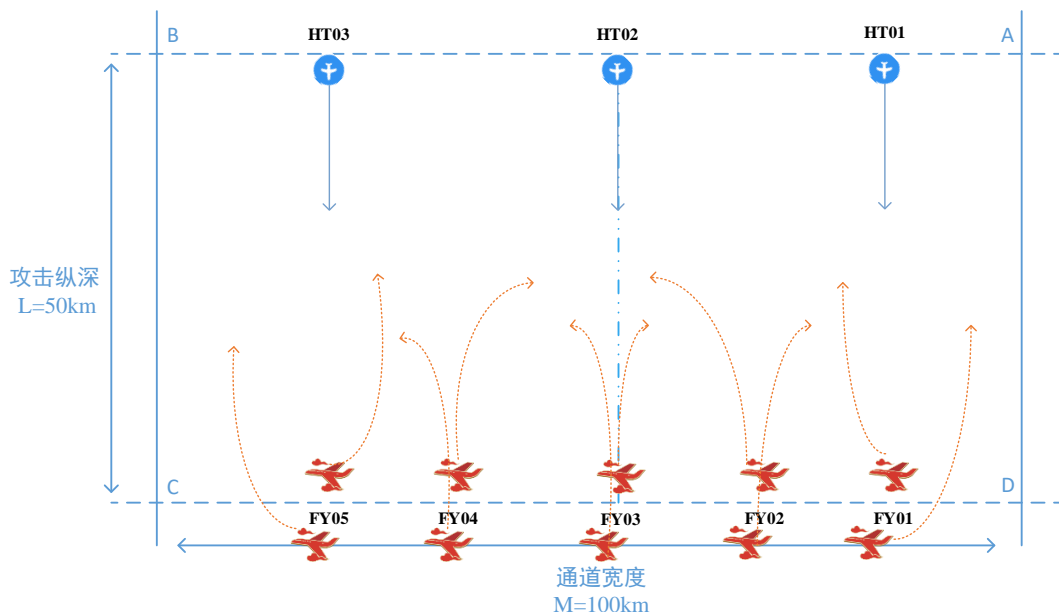


图 17 目标分配示意图

只有在最好的拦截方案下，才能学习出最优的突防策略。因此下面协同博弈模型基于此结果进行强化学习模型的建立与求解。

5.4.3 基于强化学习的协同博弈模型

本题中的红蓝双方无人机数量都有明显增加，考虑到战场平面上每一个无人机上都装有实时探测装置和数据链。假定各方无人机内部都可以实时探测到彼此的位置，并能及时做出对完成任务最有帮助的自我路线调整，那么问题转化为多智能体协同博弈对抗问题，这也更加适用于真实的智能战场环境。我们选用基于图注意力神经网络的多智能体博弈模型 G2Anet^[12]进行建模。

1、基于完全图的无人机关系构建

在本题的多智能体场景中，并不是每个无人机都能够相互作用，因此需要确定每架无人机之间的关系是否交互。那么节点之间存在一系列权重值，它们对应着无人机之间的关系。

2、基于二级注意力机制的博弈抽象 G2Net

博弈抽象的主要思想是将多智能体强化学习（马尔可夫博弈）的问题简化为一个较小的博弈，从而降低求解或学习博弈均衡策略的复杂性。而注意力机制两个主要类型：硬注意力和软注意力。前者判断两两无人机之间是否存在相互作用，它从红蓝双方一共 13 个智能个体中选择一个无人机进行判断，是不可微的，因此不能直接通过端到端的反向传播来学习注意力的权重；后者计算两架无人机之间相互作用的重要性分布，是完全可微的，因此可以通过端到端的反向传播来训练。

首先用硬注意力机制来学习能够确定无人机之间是否有交互关系的硬权重，得到任意一架无人机的关系子图 G ，在 G 中该无人机刚好和需要协同或者敌对的无人机连接。然后，使用软注意力机制来学习子图 G 中每条边的权重。

最后将上述检测机制集成到基于图神经网络的多智能体强化学习中，进行博弈抽象。

3、基于博弈抽象的演员评论家网络 GA-AC

将上述的 G2Net 和 Q 网络 AC 结合，构建出基于博弈抽象的演员评论家网络。在这个网路中，无人机之间的关系和重要性已知，在每个位置上每架无人机的奖励值或者惩罚值也能通过 AC 学习得到。关键点就在于如何将无人机间相互作用和 Q 值联系起来，从而学习到红蓝双方每架无人机在矩形区域 ABCD 中的某个位置的下一个位置，所有位置连城轨迹线所获得的收益在博弈均衡的情况下达到最大。这样可以得到每架无人机的飞行轨迹，从而确定红方的最优拦截策略和蓝方的最优突防策略。

在进行模型求解时，可以将平面进行合理的栅格化，每一个小栅格即表示无人机的位置点，这样可以大大降低计算复杂度。

因为时限原因，我们对第四题的最终策略方案只提出上述的建模构想，没能进行实际的仿真求解工作，我们会在后续研究中继续完善这项内容。

六、模型的评价

6.1 模型的优点

(1) 微分法路径规划模型使用相邻两个时刻间的位置关系式可以拟合出红、蓝双方的飞行轨迹线方程，将相遇问题转化成联立方程求解问题，并离散化平面区域以降低计算复杂度，能快速得到准确的结论。

(2) 基于 BP 神经网络的目标分配模型借助 BPN 不需先验信息的优点，有效突破题设中较多信息不明的困境，通过设定目标和约束，训练所有可能出现的样本，得到最佳的红方无人机分配方案。

(3) 问题二、三的模型有严谨的数学基础，仿真结果与理论相符合。

6.2 模型的缺点

(1) 我们不考虑红方运载机与蓝方无人机的相互影响，可能会对结果产生影响。

(2) 基于强化学习的协同博弈模型停留在理论构想阶段，尚未完成实验验证并得出预期结论。

七、模型的改进与推广

7.1 模型的改进

(1) 考虑运载机与蓝方无人机之间的约束关系，可使仿真结果更可信。

(2) 基于强化学习的协同博弈模型中，详细的参量关系尚未给出，后续可以建立严谨的理论公式，帮助完成仿真验证。

7.2 模型的推广

(1) 利用微分博弈对无人机作战的模拟场景进行兵棋推演，可以逼真地预测战场的实际作战活动，用于评估战术可行性、胜败及战争损耗，有重要的参考价值。

(2) 基于 BP 神经网络的目标分配模型同样适用于实际战场环境，训练深度 BP 神经网络可以进一步处理更复杂的战场部署问题。

参考文献

-
- [1] 田伟. 无人作战飞机航路规划研究[D]. 西北工业大学, 2007.
 - [2] 谢剑. 基于微分博弈论的多无人机追逃协同机动技术研究[D]. 2015.
 - [3] 茹常剑, 魏瑞轩, 郭庆,等. 面向无人机自主防碰撞的认知博弈制导控制[J]. 控制理论与应用, 2014, 31(011):1555-1560.
 - [4] 杨俊岭, 周宇, 王维佳, et al. 基于演化深度神经网络的无人机协同无源定位动态航迹规划[C]// cnki, 2018:26-32.
 - [5] 祁晓明, 魏瑞轩, 周凯. 面向多威胁的无人机智能目标跟随策略设计[J]. 北京航空航天大学学报, 2016, 42(04):780-788.
 - [6] SECREST. Traveling salesman problem for surveillance mission using partice swarm optimization [D]. wright-PattersonAFB:Air Force Institute of Technology,2003.
 - [7] O'Rourke K P,Bailey T G,Hill R,et al. Dynamic Routing of Unmanned Aerial Vehicles Using Reactive Tabu Search[J].Military Operations Research Journal, 2001(6):5 -30.
 - [8] Arsie A, Savla K, Frazzoli E. Efficient routing algorithmsfor multiple vehicles with no explicit communications[J].IEEE Trans on Automatic Control, 2009, 54(10):2302-2317.
 - [9] ALIGHANBARI M.Task assignment algorithms for teams of UAVs in dynamic environments[D].Cambridge:Massachusetts Institute of Technology,2004.
 - [10] NYGARD K E,CHANDLER P R,PACHTER M.Dynamic network flow optimization modles for air vehicle resourceallocation[C]//Proceedingsofthe 2001 American Control Conference,2001,3:1853-1858.
 - [11] ALVARO E G .Stability Analysis of Network-based Cooperative Resource Allocation Strategies[D].Columbus:Ohio State University,2003.
 - [12] Liu Y, Wang W, Hu Y, et al. Multi-Agent Game Abstraction via Graph Attention Neural Network[C]//AAAI. 2020: 7211-7218.